

Dynamic Dispersed Information and the Credit Spread Puzzle*

Elias Albagli
Central Bank of Chile

Christian Hellwig
Toulouse School of Economics

Aleh Tsyvinski
Yale University

December 24, 2013

Abstract

We develop a dynamic nonlinear, noisy REE model of credit risk pricing under dispersed information that can theoretically and quantitatively account for the credit spread puzzle. The first contribution is a sharp analytical characterization of the dynamic REE equilibrium and its comparative statics. Second, we show that the nonlinearity of the bond payoff in the environment with dispersed information and limits to arbitrage leads to underpricing of corporate debt and to spreads that over-state the probability of default. This underpricing is most pronounced for high investment grade, short maturity bonds. Third, we calibrate to the empirical data on the belief dispersion and show that the model generates spreads that explain between 16 to 42% of the empirical values for 4-year high investment grade, and 35 to 46% for 10-year, high investment grade bonds. These magnitudes are in line with empirical estimates linking bond spreads to empirical measures of investor disagreement, and substantially higher than most structural models of credit risk. The primary contribution of our paper in moving NREE models towards a more realistic asset pricing environment – dynamic, nonlinear, and quantitative – that holds significant promise for explaining empirical asset pricing puzzles.

*We thank Bruno Biais, Darrell Duffie, Tim McQuade, Monika Piazzesi, Guillermo Ordonez, and audiences at ESSET Gerzensee and Toulouse School of Economics for helpful comments, and Levent Guntay and Dirk Hackbarth for sharing their data with us. Hellwig gratefully acknowledges financial support from the European Research Council (starting grant agreement 263790).

1 Introduction

Noisy rational expectations equilibrium (NREE) models that follow Grossman and Stiglitz (1980)¹ are a powerful tool to analyze price discovery in financial markets with dispersed information. The core of these models is how prices aggregate dispersed information while equating demand and supply of securities. These models face three important limitations. First, they are typically static as it is challenging to analyze dynamic NREE equilibria.² Second, the payoffs are typically linear, or take on a simple parametric form, as the solution method typically relies on parametric guesses of the equilibrium price function.³ This assumption however precludes the analysis of securities with asymmetric, nonlinear payoffs, such as bonds or options. Third, the analysis is primarily qualitative and is rarely used to deliver quantitative predictions that match empirical facts.⁴ Thus, despite the importance of the framework and more than 30 years since the publication of Grossman and Stiglitz, NREE models are still far from being a mainstream tool for asset-pricing.

In this paper we propose a dynamic, non-linear generalization of NREE models, which aims to overcome these three limitations. Specifically, we develop a dynamic NREE model of corporate bond pricing with dispersed information, and argue that it delivers a novel theoretical and quantitative explanation for an important asset pricing puzzle – the credit spread puzzle.

The high levels of corporate bonds spreads relative to historical default data are difficult to reconcile with standard models pricing credit risk. Huang and Huang (2012) present a number of structural models and show that once these are calibrated using historical defaults, they all produce spreads relative to treasuries that fall short of their empirical counterparts. This shortfall is most severe for short maturity, high investment grade securities, and cannot be accounted for by standard explanations such as tax asymmetries, liquidity, and conversion options, which fail to explain the

¹See Hellwig (1980), and Diamond and Verrecchia (1981). For textbook discussions, see Veldkamp (2011) and Vives (2008).

²Important exceptions are the infinite horizon models of Wang (1993, 1994), and the multi-period model in He and Wang (1995), who study asset price dynamics and trading volume under the CARA-normal setting. Using infinite horizon models with an overlapping generation structure, Bacchetta and Van Wincoop (1995) study the connection between dispersed information and the exchange rate disconnect puzzles, while Watanabe (2008) analyzes the link between information dispersion and asset price correlation structure in the context of multiple-equilibria. Townsend (1983) analyzes higher order beliefs that emerge from forecasting the forecasts of others in the context of long-lived private information.

³Exceptions include the binary payoff model of Barlevy and Veronesi (2000, 2003), the option pricing model of Vanden (2007), and the generic extension to exponential family distributions of Breon-Drish (2012).

⁴Recent exceptions are Biais, Bossaerts and Spatt (2010), and Banerjee (2011), who test the empirical implications of dynamic REE models using stock market data. The former find that a price-contingent strategy which optimally incorporates price information outperforms a passive indexing strategy. The latter focuses on the qualitatively opposite predictions between REE and “agree to disagree” models, finding evidence in support of the REE framework.

relative spreads of corporate bonds across different credit qualities.⁵ Moreover, this puzzle is important not only for the challenge it poses on asset pricing theory, but for its potential implications on corporate investment and the volatility of real economic activity.⁶

The dynamics and the non-linearity in our model come from the nature of the asset we study – a bond pays off at a specified date in the future and can be defaulted on at any given time before maturity. We derive an analytical characterization for the prices and for the comparative statics of yield spreads with respect to credit quality, time to maturity, and the information aggregation frictions in the financial market. The combination of nonlinear bond payoffs (due to default risk), dispersed information and limits to arbitrage is essential for our theory. With dispersed information and limits to arbitrage, market-clearing causes the price to react more strongly to realizations of the fundamentals than expected payoffs. From an *ex ante* perspective, the market price then places excessive weight on tail risk. With bonds, traders are mostly concerned about default risk, so the market magnifies credit spreads relative to default risk. This amplification force is strongest for high quality and short maturity bonds.

We calibrate our model to show that the dynamic informational friction can explain a significant part of credit spreads. The performance of our model is significantly better than most structural models of credit risk, especially in pricing investment grade, short maturity bonds. The quantitative model matches existing empirical facts on the connection between dispersed beliefs and the magnitude of the credit spreads. Summarizing, we view the primary contribution of our paper as moving NREE models towards a more realistic asset pricing environment – dynamic, nonlinear, and quantitative – that holds significant promise for explaining empirical asset pricing puzzles.

Formally, we extend the non-linear noisy rational expectations equilibrium model of Albagli, Hellwig, and Tsyvinski (2012) to dynamic pricing of corporate bonds. A firm issues a bond with maturity T , which pays 1 if no default occurs before its expiration. The firm defaults if its fundamental falls below an exogenously given threshold. An investor pool is divided into informed traders who observe a noisy private signal about the firm’s fundamentals, and uninformed noise traders. Informed traders are risk neutral but face limits on their asset positions. With these assumptions, the information content of the price is particularly simple to compute, so that we can characterize prices and returns recursively without virtually any restriction on the asset’s payoff risk.

In our model, the response of bond prices to fundamental or noise trading shocks

⁵The rather large departures between the historical and predicted spreads have given rise to a growing theoretical literature enriching the set of elements that interact in different capital structure models, such as habit formation, (Chen, Collin-Dufresne and Goldstein, 2009), endogenous default (Leland and Toft, 1996; Bhamra, Kuehn and Strebulaev, 2010) and market pricing of multiple risk factors (McQuade, 2013).

⁶This point was recently stressed by Gomes and Schmid (2013), and Gourio (2013).

differs systematically from the response of expected bond payoffs to those same shocks. The bond price and the expected payoffs both incorporate the information that is conveyed through the price in equilibrium. In addition, the price must adjust to shocks in order to clear the market: if demand increases either due to better fundamentals, or due to increased noise trader demand, then the equilibrium price must also go up in order to clear the market. This market-clearing effect compounds the information conveyed through the price increase. Hence, prices react more to shocks than the expectations of fundamentals.

This stronger reaction of prices to changes in fundamentals generates spreads in excess of objective default probabilities. Intuitively, when the price signal about the bonds is low, that is, the market becomes pessimistic, the price falls more than the expected payoffs. The converse applies for high signals, with the price increasing more than the expected payoffs. Because bond payoffs are bounded on the upside and variations in payoffs are concentrated in the leftmost part of the fundamental distribution, the stronger negative reaction of prices for low fundamentals dominates the positive reaction of the prices for high fundamentals. Expected prices are then lower than expected payoffs, delivering spreads in excess of the ex-ante default probability of the security. In a dynamic context, the increase in credit spreads propagates across time, since investors anticipating a large spread (or a low bond price) in period $t+1$ reduce the price they are willing to pay for the bond, demanding a higher spread in period t . This explains why noisy information aggregation leads to an under-pricing of bonds, relative to the objective default risk.

After characterizing the equilibrium, we show several comparative statics on the effects of credit quality, bond maturity and information frictions on yield spreads. We derive several predictions, all of which are consistent with empirical evidence: First, the yield spreads exceed the underlying bond loss rates. Second, the ratio of yield spreads to loss rates is increasing in bond quality and information frictions, and ceteris paribus, information frictions have a larger effect on spreads of more highly rated securities. Intuitively, when the ex-ante default risks are small, overweighting extreme realizations has a disproportionately large impact on the market yields. When instead the ex-ante default probability is large, the overweighting of tail events has a much weaker impact on spreads. Indeed, we show that the spread ratio is unbounded as we increase credit quality arbitrarily (the default probability tends to zero, in the limit), but converges to 1 as the ex-ante default probability nears 50%, or when the informational frictions disappear. Third, the ratio between yield spreads and loss rates is decreasing in maturity. Bonds with infinitesimal maturity (i.e., default probability tending to zero) have unbounded spread ratios, but as the horizon grows large (i.e., default becomes certain, in the limit) the spread ratio converges to unity. The ability of our model to generate

large spread ratios for the highest quality bonds is particularly important, since those exhibit the largest empirical spreads relative to default rates, and therefore constitute the most significant challenge to structural bond-pricing models.⁷

We then calibrate our model to quantify how much of the observed spreads can be accounted for within our model. A key parameter in this calibration is the degree of informational frictions, which captures how much the market overweighs the tails due to noise trading and private information. We determine the range of possible values for this parameter from the measured dispersion of the analyst earnings forecasts and the volatility of firms' earnings. We then treat each bond rating/maturity as a different asset, and set the model's objective default probability to match the historical default probability for each rating/maturity pair taken from Moody's, as reported by Huang and Huang (2012). For empirically plausible levels of forecast dispersion, our model accounts for 16%-42% of the spreads on 4-year high-investment grade bonds, and 35%-46% for 10-year, high investment grade bonds. Our model is thus able to account for a fraction of the observed credit spreads that is large in comparison to the structural models reviewed by Huang and Huang (2012), and comparable to the ones delivered by the multiple risk factors model in McQuade (2013).⁸

The model also explains a sizeable fraction of the large empirical Baa-Aaa spread, one of the main puzzles of bond pricing data. Accounting for this large spread is challenging because it is unlikely to be determined by tax asymmetries, callability features, or liquidity premia, which should be roughly similar across high investment grade securities (see Chen et al., 2009). In particular, for 4 year bonds, our model explains 53% of the 103 bp spread in the data,⁹ and about 43% and 31% of the Baa-Aaa spread for 10 and 20 year bonds. These results are consistent with the findings in Guntay and Hackbarth (2010), and Buraschi, Trojani, and Vedolin (2008), two recent papers that document a significant positive relation between belief dispersion and credit spreads. Using reduced-form estimates, these papers find that forecast dispersion may statistically account for up to 1/3 of the observed level and variation in credit spreads.¹⁰

Our paper contributes to the recent literature studying the credit spread puzzle, including the papers noted above. Duffie and Lando (2001) is the closest paper to our study. They argue that imperfect observation of firm's current fundamentals generates short-term uncertainty relative to the perfect information benchmark, which increases

⁷See the discussion in Huang and Huang (2012) and McQuade (2013).

⁸For low-investment grade, the model explains an even larger fraction of the absolute spreads, with magnitudes in line with the previous results from other structural models.

⁹Spreads corresponds to those reported by Duffee (1998).

¹⁰Yu (2005) finds that firms with more opaque accounting pay larger bond spreads, particularly at short-maturities. While the mapping between our model and their empirical strategy is less clear, the evidence reported is broadly consistent with our mechanism to the extent that more noisy disclosures lead to higher levels of belief dispersion.

short-term credit spreads (through Jensen’s inequality, for concave bond payoffs). The main departures of our study from their model are twofold. First, and most importantly, we allow information to be dispersed. The resulting trading frictions then give rise to the information wedge in bond prices as a result of market-clearing forces. Second, we assume that investors disagree in equilibrium about *future* firm conditions, but current firm fundamentals are observed without noise. Hence, our finding that the model is particularly successful at explaining short-term spreads is not related to the specific modeling of short vs. long-term uncertainty, but more generally to the endogenous overweighting of tail events, which is more important for safer (i.e., shorter-term) securities.

Huang and Huang (2012) document the failure of structural models for pricing credit risk. Chen, Collin-Dufresne, and Goldstein (2009) use habit formation in preferences to explain the Baa-Aaa credit spread. Bhamra, Kuehn, and Strebulaev (2010) embed a long-run risks pricing kernel with a capital structure model featuring endogenous default, to study the term structure of Bbb-Aaa credit spreads and the equity premium in the cross-section of firms. McQuade (2013) introduces a second priced risk-factor, time-varying volatility, which together with endogenous default creates an option value component that helps explain the credit spread puzzle for different bond ratings. Gabaix (2012) and Gourio (2013) suggest a rare disasters approach. In these models, an exogenous increase in the probability of extreme events generates substantial movement in bond risk premia, without a noticeable increase in empirical default frequencies.

While these papers focus on a combination of capital structure elements and risk aversion under common information and no-arbitrage, our approach uses risk-neutral agents and exogenous default thresholds, and explores the effects of informational frictions as the only channel amplifying credit spreads.¹¹

Our paper also relates more generally to the literature on limits of arbitrage (see Gromb and Vayanos, 2010, for a recent overview). Any mispricing must result from some source of noise affecting the market, coupled with limits on the ability to exploit the resulting arbitrage opportunity. In our model, noise trading under heterogeneous information leads to systematic, predictable departures of the price from the asset’s fundamental value. Importantly, limited arbitrage together with rational, but heterogeneous beliefs lead to an *underpricing* of bonds. This contrasts with the predictions of models with heterogeneous priors and short-sales constraints,¹² in which securities

¹¹Buraschi, Trojani, and Vedolin (2011) price corporate bonds in a model where agents agree to disagree about the volatility of firms’ earnings, hence linking the stochastic discount factor to measures of economic uncertainty and disagreement.

¹²See Miller (1977), Harrison and Kreps (1978), Scheinkman and Xiong (2003), and Hong and Sraer (forthcoming).

are unambiguously overpriced due to an option value of resale. In our model, whether informational frictions and limited arbitrage lead to under- or over-pricing of securities is determined by the distribution of the underlying security's cash flow risk, and do not depend on the specific assumptions made about the nature of trading constraints.

Section 2 presents the main empirical facts on the credit spread puzzle and reviews them from the perspective of no-arbitrage models. Section 3 introduces our model and derives the equilibrium characterization. Section 4 derives comparative statics results for credit spreads for certain special cases. Section 5 presents our calibration exercise and quantitative results. Section 6 concludes. All proofs are in the appendix.

2 The Credit Spread Puzzle

2.1 Empirical Facts

In this sub-section, we document the main stylized facts regarding the credit spread puzzle. For credit ratings from Aaa to B and bond maturities of 4, 10, and 20 years, table 1 reports average observed yield spreads, cumulative default probabilities, and average annualized loss rates. We impute these statistics assuming (as in Huang and Huang, 2012) that in a default, creditors recover on average 51% of the face value of the debt. Finally, the table reports the *spreads ratio* of the yield spreads observed in the market to the annualized loss rates. The difference between yield spreads and loss rates measures the excess return on bonds, while the spreads ratio summarizes by how much the market over-estimates actual default risks. This measure also plays a central role in our theoretical analysis.

The following two observations are particularly striking:

Stylized Fact I: *Annualized loss rates cannot account for more than 15% of the average level of credit spreads on investment-grade bonds.*

Stylized Fact II: *The fraction of credit spreads explained by default risk is much lower for safer bonds and/or for bonds with shorter maturities.*

Table 1: Historical Default rates, Yield Spreads, and Spread Ratios

	Average Yield Spread (bps) ¹			Cumulative Default Rates ² (%)		
	4 yr	10 yr	20 yr	4 yr	10 yr	20 yr
Aaa	46	47	59	0.04	0.77	2.05
Aa	56	69	87	0.23	0.99	2.32
A	87	96	117	0.35	1.55	4.45
Baa	149	150	198	1.24	4.39	10.89
Ba	310	310		8.51	20.63	36.99
B	470	470		23.32	43.91	52.90

	Annualized loss rates (bps) ³			Spread Ratio ⁴		
	4 yr	10 yr	20 yr	4 yr	10 yr	20 yr
Aaa	0.5	3.8	5.0	93.87	12.43	11.69
Aa	2.8	4.9	5.7	19.86	14.19	15.22
A	4.3	7.6	11.0	20.27	12.59	10.61
Baa	15.2	21.7	27.4	9.78	6.90	7.22
Ba	106.5	106.6	100.0	2.91	2.91	
B	303.4	242.3	150.0	1.55	1.94	

¹ Yield Spreads for Aaa-Baa are from Duffee (1998); Ba and B from Caouette, Altman, and Narayanan (1998).

² Cumulative default rates are from Moody's report by Keenan, Shtogrin, and Sobehart (1999), covering the period 1970-1998.

³ Annualized loss rates are computed using the average recovery rate of 51%, and are calculated as $-(1/T)\ln(1-0.49 \cdot CDR)$, where CDR is the cumulative default rate at maturity T .

⁴ The spread ratio is the ratio between the average yield spread and the annualized loss rates.

Table 2: Observed vs. calculated credit spreads

	Average Yield Spread* (bps)		Calculated Yield Spreads		Fraction explained (%)	
	4 yr	10 yr	4 yr	10 yr	4 yr	10 yr
Aaa	55	63	1.1	10.0	2.0	15.9
Aa	65	91	6.0	14.2	9.2	15.6
A	96	123	9.9	23.3	10.3	18.9
Baa	158	194	32.0	56.5	20.3	29.1
Ba	320	320	172.3	192.3	53.8	60.1
B	470	470	445.7	387.8	94.8	82.5

¹ Average yield spreads and calculated yield spreads are taken from Table 1 of Huang and Huang (2012).

Table 1 made no attempt to impute risk premia on the credit default risk. Huang and Huang (2012) study how much of the excess return can be attributed to credit risk. They test a large class of structural models calibrated to be consistent with data on the historical default loss experience and equity risk premia. They find that for the investment-grade bonds of all maturities, credit risk accounts for only a small fraction of

the observed yield spreads, on average about 20% for a typical model. Table 2 reports the spreads predicted by the baseline structural model in Huang and Huang (2012), both in absolute levels, and as a fraction of the observed spreads. As an example, for the Baa bonds with maturity of 4 years they find that the (baseline) model predicts the spread of only 32 bp, significantly below the observed spread of 158 basis points. The puzzle is much more significant for the bonds of higher ratings. For junk bonds the credit risk accounts for a much larger fraction of the observed yield spreads.

These columns summarize the credit spread puzzle: while risk premia on default risk explain a large fraction of spreads for speculative grade debt, they account for only a small proportion of the spreads for high investment grade corporate bonds, specially for the short maturity bonds. The results reported from Huang and Huang (2012) have been confirmed by several other empirical studies. For example, Elton, Gruber, Agrawal and Mann (2001) find that, in a risk-neutral setting, the expected default losses can account for no more than 25% of the corporate spreads. Collin-Dufresne, Goldstein and Martin (2001) study the determinants of the credit spread changes (rather than levels of the credit spreads) for corporate bonds. They find that various proxies for changes in the default probability and the recovery rates can account only for about 25% of the observed credit spread changes.

The third stylized fact links credit spreads to measures of investor disagreement.

Stylized Fact III: *Credit spreads increase with forecast disagreement, accounting for as much as 1/4 to 1/3 of the observed level and variation of spreads.*

This fact is established independently by Guntay and Hackbarth (2010) and Buraschi, Trojani and Vedolin (2008). Proxying for disagreement with analyst forecast dispersion of forthcoming quarterly earnings per share, Guntay and Hackbarth document that forecast dispersion is a highly significant and economically important predictor of spreads in univariate, cross sectional regressions, explaining about 23% of the cross-sectional variation in credit spreads. In panel-data regressions, they find that forecast dispersion predicts an average sample spread of about 14 bp, which is about 14% of the mean credit spread in their sample. Moreover, a +1 standard deviation increase in forecast dispersion is associated with an increase in the spread of another 14 basis points.¹³

Buraschi, Trojani and Vedolin (2008) use dynamic factor analysis to construct a series of firm earnings disagreement that has both a systematic and idiosyncratic component. They show that both disagreement proxies have an unambiguously positive impact on credit spreads. More specifically, a one standard deviation increase in firm-

¹³Interestingly, these results for bond returns are the opposite of what is found for equity returns, which are negatively related to disagreement measures (see Diether, Malloy, and Scherbina 2002, and Johnson 2004).

specific disagreement increases credit spreads by approximately 18 basis points, which is more than one third the credit spread sample standard deviation in their data. Similarly, a one standard deviation increase in systematic disagreement increases the average credit spread by about 10 basis points. Both disagreement proxies together generate about 44 bps, which constitute approximately about 30% of their sample mean of 142 bps across all bond categories.

2.2 Credit spreads from a no-arbitrage perspective

In this subsection we discuss the challenge of accounting for the observed credit spreads in risk-based no-arbitrage models. For simplicity, we focus on the return on bonds that are held to maturity. We conduct two simple exercises. First, we bound the worst-case scenarios for bond returns that the market must consider possible (regardless of the pricing kernel used to price the bonds), and show that this worst-case scenario is far outside the range of realizations, in terms of default losses, that are observed in the last 30 years of data. Second, we compute the Sharpe ratios for corporate bonds held to maturity from the distribution of default risks and the observed credit spreads, and argue that the stochastic discount factor required to price these bonds must be an order of magnitude more volatile than the Sharpe Ratio bound derived from equity returns.

Let m denote a stochastic aggregate state at maturity (w.l.o.g., this state is identified with the stochastic discount factor). Let $R_i(m)$ denote the average return on a portfolio of bonds with characteristic i , when aggregate state m is realized. $R_i(m)$ can be written as $R_i(m) = \bar{R}_i \cdot (1 - L_i(m))$, where \bar{R}_i is the initial yield (hence $1/\bar{R}_i$ the initial price), and $L_i(m)$ the aggregate loss rate on bonds with characteristic i in state m . This loss rate is the fraction of bonds that default, i.e. pay less than its face value, times one minus the recovery value of defaulting bonds. The bond returns must satisfy the moment condition $\mathbb{E}(R_i(m) \cdot m) = 1$. The annualized yield spread s_i of a bond with T years to maturity is $s_i = 1/T \cdot \log(\bar{R}_i/R_f)$, where $R_f = 1/\mathbb{E}(m)$ is the risk-free return. The annualized loss rate is $d_i = -1/T \cdot \log(1 - \mathbb{E}(L_i(m)))$. The difference $s_i - d_i$ then measures the annualized excess return of the bond.

Table 3 is constructed from the yearly default rate statistics in the Moody's report of Keenan, Shtogrin, and Sobehart (1999), covering the period between 1970-1998. We choose this time interval throughout the paper to facilitate the comparison with the papers cited above (which use a similar time period), and because the results and main conclusions are not sensitive to including more recent data.¹⁴ In particular, we report summary statistics on the time variation of cumulative loss rates at the 4 and

¹⁴Chen et al. (2009) report corresponding numbers on the average and variation of the cumulative default rate of Baa bonds of 1.55% (average), 0% (min), 3.88% (max) and 1.04% (standard deviation), for the period of 1970-2001.

10-year horizon for each rating categories, as well as for each broad bond category (i.e., investment vs. speculative grade). Besides the time-series average of cumulative loss rates, we report information on both the worst realizations and the best realizations from this sample of roughly 30 yearly vintages, the sample standard deviation, and the sample Sharpe ratio.

Table 3: Cumulative loss rate statistics, 1970-1999

4 yr	Aaa	Aa	A	Baa	Ba	B	Inv. grade	Spec. grade	All
Average (%): E (Li (m))	0.02	0.11	0.17	0.61	4.26	12.13	0.28	7.34	2.14
Max (%): Sup _m (Li (m))	0.59	0.43	1.12	1.91	10.76	21.72	0.79	15.04	5.31
Min (%): Inf _m (Li (m))	0.00	0.00	0.00	0.00	0.86	1.86	0.00	1.71	0.36
St. Dev. (%): σ (Li (m))	0.12	0.16	0.29	0.54	3.23	5.78	0.24	4.27	1.49
Sharpe Ratio: E (Li (m)) / σ (Li (m))	0.17	0.70	0.59	1.12	1.32	2.10	1.18	1.72	1.44

10 yr	Aaa	Aa	A	Baa	Ba	B	Inv. grade	Spec. grade	All
Average (%): E (Li (m))	0.38	0.49	0.76	2.17	10.66	24.23	1.12	14.94	4.28
Max (%): Sup _m (Li (m))	1.58	1.24	2.32	4.71	19.33	34.94	2.53	23.80	7.06
Min (%): Inf _m (Li (m))	0.00	0.00	0.00	1.44	2.94	3.70	0.63	4.13	1.53
St. Dev. (%): σ (Li (m))	0.53	0.38	0.71	1.01	6.10	10.89	0.55	7.52	2.14
Sharpe Ratio: E (Li (m)) / σ (Li (m))	0.71	1.28	1.08	2.16	1.75	2.22	2.03	1.99	2.00

Bounding the worst case scenario: Restate the moment condition as

$\mathbb{E}(R_i(m)/R_f \cdot m/\mathbb{E}(m)) = 1$. Since $R_i(m)/R_f = \bar{R}_i/R_f \cdot (1 - L_i(m))$, one obtains

$$\sup_m L_i(m) \geq \mathbb{E}\left(L_i(m) \cdot \frac{m}{\mathbb{E}(m)}\right) = \frac{\bar{R}_i - R_f}{\bar{R}_i} = 1 - e^{-s_i T} \approx s_i T.$$

Simply put, this condition says that there must be at least some chance that default losses more than offset the yield spreads. Otherwise the bond would offer a return in

excess of the risk-free rate with probability one, and hence a sure arbitrage opportunity. This places an obvious lower bound on the worst case losses that the market must consider possible.

Over the sample considered (1970-1998), the worst loss realization for all 4 year corporate bonds was 5.31%, originating from the 1989 vintage. Looking at each rating category, the worst loss rates were 0.59%, 0.43%, 1.12%, 1.91%, 10.76%, and 21.72% for Aaa through B categories. In comparison, the yield spreads reported in table 1 suggest that the minimal worst case realization at the 4 year horizon that is required to be consistent with a stochastic discount factor are $1 - e^{-0.0046 \cdot 4}$, or 1.82% for Aaa-rated bonds, and 2.22%, 3.42%, 5.79%, 11.66%, and 17.14% for Aa through B-rated bonds. Similarly at the 10-year horizon, the worst loss rate realizations were 1.58%, 1.24%, 2.32%, 4.71%, 19.33%, and 34.94% for the Aaa through B categories. In comparison, the theoretical bounds are 4.59%, 6.67%, 9.15%, 13.93%, 26.66%, and 37.50% for Aaa through B categories.

Hence, for investment grade bonds (Aaa-Baa), the worst loss realizations that observed fall systematically short of the bound imputed from observed yields. In other words, the excess return on investment grade bonds was sufficient to fully cover the default losses incurred in even the worst years over 30 years of data. In fact, under a no-arbitrage view of the world, the market must put significant weight on outcomes that are between 3 and 5.15 times worse than the worst realizations over this sample period for 4 year bonds, and between 2.9 and 5.38 times for 10 year bonds.¹⁵

For speculative grade instruments on the other hand, theoretical bounds are much closer to the empirical worse-case realizations, and in some cases the empirical worst cases exceed the theoretical bounds: for 4 year bonds, compare the 10.76% empirical worse realization with the 11.66% theoretical worse outcome for Ba rated bonds, and 21.72% vs. 17.14% for B rated bonds. For 10 year bonds, these figures are 19.33% vs 26.66% for Ba, and 34.94% vs. 37.5% for B rated bonds.

This simple diagnostic illustrates well the challenge of accounting for observed investment-grade bond yields with risk-based no-arbitrage models. On the other hand, speculative grade bond yields are much easier to account for with stochastic discount factor models, a finding that is fully consistent with the existing literature.

Sharpe Ratios for Corporate Bonds: Next, we compute a lower bound on the variance of m . Substitute $\mathbb{E}(R_i(m)) = R_f \cdot e^{(s_i - d_i)T}$, $\mathbb{E}(L_i(m)) = 1 - e^{-d_i T}$, and

¹⁵For example, the worst cumulative loss rate on 4 year Aa bonds was 0.43%, while the corresponding bound suggests a minimal worst case scenario of 2.22%, giving a ratio of 5.15. For 10 year Aa, the worst cumulative loss rate was 1.24%, while the corresponding 10-year bound suggests a minimal worst case scenario of 6.67%, giving a ratio of 5.38.

$\sigma(R_i(m)) = \bar{R}_i \cdot \sigma(L_i(m))$, to rewrite the Hansen-Jaganathan inequality as follows:

$$\frac{\sigma(m)}{\mathbb{E}(m)} \geq \frac{\mathbb{E}(R_i(m)) - R_f}{\sigma(R_i(m))} = \frac{e^{(s_i-d_i)T} - 1}{e^{s_i T} - e^{(s_i-d_i)T}} \frac{\mathbb{E}(L_i(m))}{\sigma(L_i(m))} \approx \frac{s_i - d_i}{d_i} \frac{\mathbb{E}(L_i(m))}{\sigma(L_i(m))}$$

Thus, the RHS can be written as the product of the Sharpe ratio for the bonds' loss rate, and a term that just depends on the ratio of yield spreads to loss rates. In other words, this is a measure of the excess return per "unit" of default risk. The last row of Table 3 reports Sharpe ratios for the loss rates on Aaa through B-rated bonds at 4- and 10-year horizon, while Table 1 provide direct measures of yield spreads and default loss rates to compute the ratio $(s_i - d_i)/d_i$. For 4-year bonds, we obtain the following lower bounds on $\sigma(m)/\mathbb{E}(m)$: 15.78, 13.21, 11.37, 9.83, 2.52, and 1.15 for Aaa through B-rated bonds. Likewise, for 10-year bonds, these numbers are 8.12, 16.88, 12.52, 12.74, 3.34, and 2.09 for Aaa through B categories. Even accounting for the difference in time horizon, these bounds are an order of magnitude larger than the bounds obtained from equity returns.

Alternatively, we can bound $\mathbb{E}(L_i(m))/\sigma(L_i(m))$ using only support restrictions for $L_i(m)$, without relying on the empirical Sharpe Ratios. Let $\bar{L}_i = \sup_m L_i(m) \leq 1$ denote the worst-case scenario for aggregate loss rates that the market considers possible. Since $\sigma^2(L_i(m)) \leq \mathbb{E}(L_i(m))(\bar{L}_i - \mathbb{E}(L_i(m)))$, we have $\mathbb{E}(L_i(m))/\sigma(L_i(m)) \geq 1/\sqrt{\bar{L}_i/\mathbb{E}(L_i(m)) - 1}$. Using the numbers from table 3, this bounds the Sharpe ratios for the bonds' loss rate in the range from 0.19 for Aaa to 1.12 for B-rated, 4-year bonds, and in the range from 0.56 for Aaa to 1.5 for B-rated, 10-year bonds, roughly 1/2 to 1 times the empirical measure for $\mathbb{E}(L_i(m))/\sigma(L_i(m))$. Although lower than the actual Sharpe ratios on loss rates, this still suggests that the stochastic discount factor must be far more volatile than what is implied by equity returns, due to the very large difference between yield spreads and default rates.

Now, what would such a stochastic discount factor have to look like? It must attach a very high relative value to consumption in certain highly unlikely states with a large incidence of defaults, much in the spirit of the rare-disasters literature.¹⁶ In fact, Gabaix (2012) suggests a rare disasters approach as a possible resolution of the credit spread puzzle. However, default rates in these states must be much worse than any default rates observed in recent memory, and the marginal value of cash flows in the disaster state an order of magnitude higher than on average.¹⁷ Any attempt to account for

¹⁶For any given support $[\underline{m}, \bar{m}]$ and expectation $\mathbb{E}(m)$, $\sigma(m)$ is maximal if m places all its weight on the extreme realizations of m . What's more, since $\sigma^2(m) \leq (\mathbb{E}(m) - \underline{m})(\bar{m} - \mathbb{E}(m))$, the ratio $\bar{m}/\mathbb{E}(m)$ is bounded by $\frac{\bar{m}}{\mathbb{E}(m)} \geq 1 + \left(\frac{\sigma(m)}{\mathbb{E}(m)}\right)^2$. Using the computed H-J bounds, this suggests a very large gap between marginal values in the disaster state, relative to the average.

¹⁷Giesecke, Longstaff, Schaefer and Strebulaev (2011) collect data on aggregate US non-financial issuers for the 1866-2008 period, and find that indeed several of the worse default episodes occur in the 19th century,

observed credit spreads in the context of a no-arbitrage model thus rests on premise that observed bond yields account for aggregate default risk not observed in the data, and the utility costs of such default risk (in terms of fluctuations in the marginal utility of the representative investor) must be exceedingly large.

To conclude, these simple calculations offer perhaps the most direct illustration yet that observed credit spreads are difficult to reconcile with a no-arbitrage model of default risk and bond returns. This further motivates the alternative route that we explore here, which introduces noisy information aggregation and limits to arbitrage to account for the observed patterns of credit spreads.

3 The model

In this section, we extend the noisy rational expectations model of asset prices of Albagli, Hellwig, and Tsyvinski (2012) to dynamic pricing of bonds.

3.1 Model environment

Fundamentals and bond payoffs: Let θ_t denote the fundamental of a firm at period t . The fundamental follows an AR(1) process,

$$\theta_{t+1} = \rho\theta_t + \varepsilon_{t+1}, \tag{1}$$

where $\varepsilon_{t+1} \sim \mathcal{N}(0, \sigma_\theta^2)$ and $\rho \in (0, 1)$. The firm enters into default the first time that the fundamental θ_t falls below an exogenous threshold $\underline{\theta}$. Default is an absorbing state. The fundamental thus summarizes the firms' financial health, or its distance to bankruptcy. We treat the fundamental process as exogenous and focus on the bond pricing implications of defaults. We consider a zero coupon bond that pays 1 at maturity T , if and only if $\theta_t \geq \underline{\theta}$ for all $t \leq T$, and $c \in (0, 1)$ if it defaults before maturity. The parameter c represents the recovery rate for bondholders in case of a default. The supply of bonds is normalized to a unit measure.

Investors: There is a unit measure of informed investors who are risk-neutral and live for one period. They receive a noisy private signal $x_t \sim \mathcal{N}(\theta_{t+1}, \beta^{-1})$ before deciding whether to purchase the bond in period t to re-sell it in period $t + 1$. Their bond holdings are restricted to the unit interval.¹⁸ In addition, traders perfectly observe

mostly related to the railroad crisis of the 1870's.

¹⁸The role of the risk-neutrality and position bound assumptions for the equilibrium characterization as well as generalizations of such environment are discussed at length in Albagli, Hellwig and Tsyvinski (2012).

the current fundamental θ_t at the beginning of each period. The dispersion of beliefs thus resolves at the end of each period. In addition, there are noise traders who buy a fraction $\Phi(u_t)$ of the bonds, where $u_t \sim \mathcal{N}(0, \sigma_u^2)$ is iid over time, and $\Phi(\cdot)$ denotes the cdf of a standard normal distribution.

Trading Environment and Equilibrium: In each period, informed investors submit price-contingent orders to purchase the available bonds. Noise traders bid for a fraction $\Phi(u_t)$ of the bonds. The market-clearing price P_t is then selected so that the total demand by informed and noise traders equals the available supply of 1.

We focus on a Recursive Bayesian Equilibrium, in which informed investors only condition on the current $\{\theta_t, P_t\}$, and P_t is a function only of θ_{t+1} , u_t , and θ_t and time to maturity τ . Given the knowledge of θ_t , the past history of $\{\theta_{t-s}, P_{t-s}\}_{s=1}^t$ contains no further information on current and future prices and expected payoffs, and if informed traders only condition on $\{\theta_t, P_t\}$ and x_t , then the market-clearing price must be a function of θ_{t+1} , u_t , and θ_t . A Recursive Bayesian Equilibrium consists of (i) a bidding schedule $a_i(x, \theta, P) \in [0, 1]$, that is optimal given informed traders' beliefs $H(\cdot|x, \theta, P)$, (ii) informed traders' beliefs $H(\cdot|x, \theta, P)$ which are consistent with Bayes' Rule, and (iii) a price function $P(\tau, \theta, \theta_{+1}, u)$, such that the market clears for all (θ, θ_{+1}, u) .

3.2 Recursive equilibrium characterization

Let $p(\tau - 1, \theta_{+1})$ denote the expected value of a bond with time to maturity $\tau - 1$, as a function of the next period fundamental θ_{+1} . This value is characterized recursively. To initiate the recursion, suppose that $p(\tau - 1, \theta_{+1})$ is non-decreasing in θ_{+1} , everywhere, and strictly increasing somewhere. With risk neutrality, position bounds and log-concavity of the private signals, the informed traders' demand is characterized by a threshold function. That is, the informed trader submits an order of 1, whenever his private signal $x \geq \hat{x}(\theta, P)$, and 0 otherwise. The indifference condition for the signal threshold $\hat{x}(\theta, P)$ is

$$P = \mathbb{E}(p(\tau - 1, \theta_{+1}) | \hat{x}(\theta, P), \theta, P). \quad (2)$$

Therefore, the informed traders' demand for the bond is $1 - \Phi(\sqrt{\beta}(\hat{x}(\theta, P) - \theta_{+1}))$, and the market clears if and only if

$$1 - \Phi\left(\sqrt{\beta}(\hat{x}(\theta, P) - \theta_{+1})\right) + \Phi(u) = 1,$$

or if and only if $\hat{x}(\theta, P) = z \equiv \theta_{+1} + 1/\sqrt{\beta} \cdot u$. Moreover, $\hat{x}(\theta, P)$ must be invertible w.r.t. P , otherwise market clearing must be violated for some realizations of z . The random variable z thus fully summarizes the information conveyed through the market price P . Using this fact and $\hat{x}(\theta, P) = z$ (by market-clearing), we obtain P as a function

of z , θ , and time to maturity τ , for $\theta_t \geq \underline{\theta}$:

$$P(\tau, \theta, z) = \mathbb{E}(p(\tau - 1, \theta_{+1}) | x = z, \theta, z).$$

The expected value $p(\tau - 1, \theta_{+1})$ is equal to c in case of a default, and equal to the expected bond price next period, otherwise:

$$p(\tau - 1, \theta_{+1}) = \begin{cases} \mathbb{E}(P(\tau - 1, \theta_{+1}, z') | \theta_{+1}) & \text{if } \theta_{+1} \geq \underline{\theta}, \\ c & \text{if } \theta_{+1} < \underline{\theta}. \end{cases} \quad (3)$$

Combining equations (2) and (3), the following proposition summarizes the equilibrium bond prices and expected cash values by means of a simple recursion for $p(\tau, \theta)$. The proof completes the characterization by showing that the sequences of functions $p(\tau - 1, \theta)$ is indeed monotonic in θ so that the induction step is valid.

Proposition 1. *The equilibrium price for a τ -period bond is characterized by*

$$P(\tau, \theta, z) = \mathbb{E}(p(\tau - 1, \theta_{+1}) | x = z, \theta, z), \quad (4)$$

where the expected cash value $p(\tau, \theta)$ is given recursively by

$$p(\tau, \theta) = \begin{cases} \mathbb{E}(\mathbb{E}(p(\tau - 1, \theta_{+1}) | x = z, \theta, z) | \theta) & \text{if } \theta \geq \underline{\theta} \\ c & \text{if } \theta < \underline{\theta} \end{cases} \quad (5)$$

with $p(0, \theta) = 1$ iff $\theta \geq \underline{\theta}$ and $p(0, \theta) = c$ otherwise.

In our next step we simplify the recursion for $p(\tau, \theta)$ to provide insight into how the information friction affects bond prices. In particular, notice that the recursion compounds two normal posterior expectations. Using straightforward algebra, we show the following lemma:

Lemma 1. *The expected cash value is recursively given by*

$$p(\tau, \theta) = \begin{cases} \int p(\tau - 1, \theta_{+1}) d\Phi\left(\frac{\theta_{+1} - \rho\theta}{\sigma_P}\right) & \text{if } \theta \geq \underline{\theta}, \\ c & \text{if } \theta < \underline{\theta}, \end{cases} \quad (6)$$

where $\sigma_P^2 = \sigma_\theta^2 + (1 + \sigma_u^2)D$, and

$$D = \beta / (1/\sigma_\theta^2 + \beta + \beta/\sigma_u^2)^2. \quad (7)$$

Lemma 1 states that expected bond prices are characterized recursively by a transition probability function for the sequence of fundamentals $\{\theta_t\}$. Notice that the implied variance used in this computation equals σ_P^2 , which is strictly larger than σ_θ^2 . In other words, the recursion defining bond prices assigns a larger weight to tail realizations than the objective distribution with variance σ_θ^2 . This market-implied variance of fundamentals, $\sigma_P^2 = \sigma_\theta^2 + (1 + \sigma_u^2) D$ depends on the variance of fundamental shocks σ_θ^2 , on the variance of noise trading, σ_u^2 , and on D , which measures the dispersion of traders' forecasts of the fundamental θ .

3.3 Prices, expected payoffs, and bond spreads

The expected payoff value of a bond with maturity τ and current fundamental θ , $v(\tau, \theta)$, is defined recursively as

$$v(\tau, \theta) = \begin{cases} \int v(\tau - 1, \theta_{+1}) d\Phi\left(\frac{\theta_{+1} - \rho\theta}{\sigma_\theta}\right) & \text{if } \theta \geq \underline{\theta}, \\ c & \text{if } \theta < \underline{\theta}. \end{cases} \quad (8)$$

Conditional on market information z , the expected market value of the bond in the following period is $\tilde{V}(\tau, \theta, z) = \mathbb{E}(p(\tau - 1, \theta_{+1}) | \theta, z)$, while the expected dividend value, if held to maturity is $V(\tau, \theta, z) = \mathbb{E}(v(\tau - 1, \theta_{+1}) | \theta, z)$.

To understand the forces at work in determining the magnitude of credit spreads, it is convenient to decompose the difference between the price and the expected payoff value of the bond as follows:

$$P(\tau, \theta, z) - V(\tau, \theta, z) = P(\tau, \theta, z) - \tilde{V}(\tau, \theta, z) + \tilde{V}(\tau, \theta, z) - V(\tau, \theta, z).$$

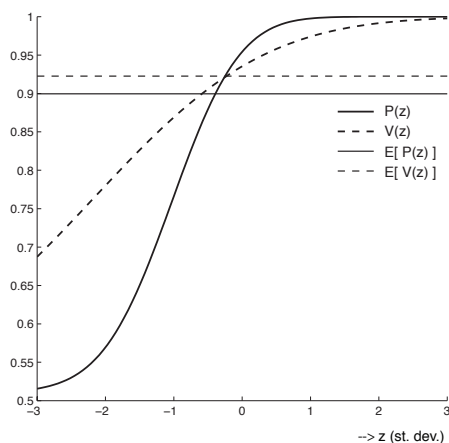
The first component is the difference between the price ($P(\tau, \theta, z)$) and the expected market value of the bond next period ($\tilde{V}(\tau, \theta, z)$). While both are expectations about the same underlying object (the expected future market value of the bond), they condition on different information, and hence differ in the weight attributed to the market signal z . The second component is the difference between the expected market value, and the expected bond payoff ($V(\tau, \theta, z)$). Both these expectations are taken under the same information (and hence weight the market signal z equally), but constitute forecasts about different statistical objects which arise in a multi-period trading environment.

To understand the first component, note that the price in (4) puts a higher weight on the signal z , relative to the expected market value of the bond $\tilde{V}(\tau, \theta, z)$. The price and the expected bond value both incorporate the information that is conveyed

through the price (or its sufficient statistic z) in equilibrium. In addition, the price must adjust to the fundamental shocks to satisfy the market-clearing condition. Even without considering the inference drawn from the price, an increase in demand resulting from a more favorable realization of the payoff fundamental or an increase in the noise traders' demand must be met by an increase in the equilibrium price in order to clear the market. This direct market-clearing effect is present only in the determination of the price but not in the expected future bond value.

Figure 1 illustrates this channel. We plot the price and the expected payoff value as a function of the market signal, for the case of a one-period bond. At the end of year 1, the bond pays contingent on the realization of θ_1 . At the market stage (period 0), prices and expected payoffs condition on the prior, and the market signal $z = \theta_1 + u/\sqrt{\beta}$. Prices $P(z)$ react stronger than the expected payoffs to the realization of the information about the fundamental. For positive news about the fundamental ($z > 0$; that is, the signal z is above the unconditional mean), the price is above the expected payoffs. The reverse is true for the negative news ($z < 0$).

Figure 1: Price vs. expected payoff, one-year bond



The key aspect to highlight from the figure is that the *magnitude* of the reaction of price and the expected payoff to the information about the fundamentals depends on the sign of z . Since the risk of the bond is concentrated on the downside – relative to its mean payoff, losses in extreme negative states are larger than gains in extreme positive states – the difference between $P(z)$ and $V(z)$ is small on the upside (when it is positive), but is large (in absolute magnitude) on the downside (when it is negative). Integrating over all of the realizations of the signal z , it is then evident that the expected prices are lower than expected cash flows.¹⁹

¹⁹Since the bond matures at the end of the first period, there is no distinction between the expected future

the expected bond value in (8) which uses σ_θ as the volatility of fundamentals and the recursion for the price equation in (6) which uses $\sigma_P > \sigma_\theta$ for the volatility of the fundamental process. From an ex ante perspective, the fact that the market price reacts more strongly to z in future periods leads to market-implied beliefs about $\{\theta_t\}$ which associate the same degree of persistence, but a higher variance for the innovations, relative to the fundamental distribution. This stronger reaction to z in future periods propagates to current bond prices through the recursion equation. Due to expected future informational frictions, the current bond price places a larger weight on the risk of future tail realizations than would be implied by the objective distribution. Since the tail risk in bond returns comes from default, current bond prices thus over-weigh future default risk. This is the channel through which information frictions propagate across time. In short, the informational frictions that cause the price to overreact to market signals, combined with the shape of bond cash flow risks, leads to a systematic underpricing of bonds, or spreads that appear high relative to the objective default probabilities.

We now transform prices and default probabilities into yields and default rates exactly as in section 2. Since the recovery rate is independent of realized states or time to maturity, we can write the expected price and expected payoffs simply as functions of the market-implied and objective cumulative default probabilities $\hat{\Pi}(\tau, \theta)$ and $\Pi(\tau, \theta)$:

$$p(\tau, \theta) = 1 - (1 - c) \hat{\Pi}(\tau, \theta) \quad \text{and} \quad v(\tau, \theta) = 1 - (1 - c) \Pi(\tau, \theta),$$

where

$$\Pi(\tau, \theta) = \Phi\left(\frac{\theta - \rho\theta}{\sigma_\theta}\right) + \int_{\underline{\theta}}^{\infty} \Pi(\tau - 1, \theta_{+1}) d\Phi\left(\frac{\theta_{+1} - \rho\theta}{\sigma_\theta}\right), \quad (9)$$

$$\hat{\Pi}(\tau, \theta) = \Phi\left(\frac{\theta - \rho\theta}{\sigma_P}\right) + \int_{\underline{\theta}}^{\infty} \hat{\Pi}(\tau - 1, \theta_{+1}) d\Phi\left(\frac{\theta_{+1} - \rho\theta}{\sigma_P}\right), \quad (10)$$

with $\Pi(0, \theta) = \hat{\Pi}(0, \theta) = 0$, for $\theta > \underline{\theta}$. The market-implied yield-spread is then given by

$$s(\tau, \theta) = -\frac{1}{\tau} \log(p(\tau, \theta)) = -\frac{1}{\tau} \log\left(1 - (1 - c) \hat{\Pi}(\tau, \theta)\right)$$

and the expected default rate is

$$d(\tau, \theta) = -\frac{1}{\tau} \log(v(\tau, \theta)) = -\frac{1}{\tau} \log\left(1 - (1 - c) \Pi(\tau, \theta)\right)$$

The yield spread $s(\tau, \theta)$ and default rates $d(\tau, \theta)$ are thus the model-implied counterparts to the empirical measures reported in table 1. In the model, they highlight the contribution of informational frictions to generate spreads in excess of those which will

obtain in a risk-neutral model with no information heterogeneity, and therefore cleanly pin down the contribution of our mechanism. If the objective default probabilities are small, we approximate the spreads as

$$\begin{aligned} s(\tau, \theta) &\approx -\frac{1}{\tau} (1 - c) \hat{\Pi}(\tau, \theta), \\ d(\tau, \theta) &\approx -\frac{1}{\tau} (1 - c) \Pi(\tau, \theta), \end{aligned}$$

and therefore the ratio of the spreads $s(\tau, \theta)/d(\tau, \theta)$ is simply given by the ratio of default probabilities $\hat{\Pi}(\tau, \theta)/\Pi(\tau, \theta)$. This ratio is independent of the recovery rate c and depends on the bond maturity only through its effect on default probability.

4 Comparative Statics

This section develops comparative statics results that emerge from our information-based theory of credit spreads. We focus on the case with small default risks, in which the spread ratio is equal to the ratio of default probabilities.

4.1 The case without re-trading ($\tau = 1$)

For one-period bonds, it is straightforward to fully derive the pricing implications of informational frictions. In this case, probabilities are given by

$$\hat{\Pi}(1, \theta) = \Phi\left(\frac{\underline{\theta} - \rho\theta}{\sigma_P}\right) \text{ and } \Pi(1, \theta) = \Phi\left(\frac{\underline{\theta} - \rho\theta}{\sigma_\theta}\right).$$

Define the variables $\xi = \sigma_\theta/\sigma_P$ and $v = (\underline{\theta} - \rho\theta)/\sigma_\theta$. The variable v measures the objective default risk associated with the bond, and therefore summarizes the credit quality (a higher v corresponds to a more distressed bond). Notice also that $\underline{\theta} - \rho\theta < 0$, for all $\theta > \underline{\theta}$, and hence $v < 0$, i.e. the ex-ante default probability is always less than $1/2$, if the current fundamental precludes an immediate default. With this parametrization, $\Pi(1, \theta) = \Phi(v)$ and $\hat{\Pi}(1, \theta) = \Phi(\xi v)$. The parameter ξ measures the severity of information aggregation frictions, and thus summarizes how significant the departure from no-arbitrage pricing is. When $\xi = 1$, the market-implied and objective variance of fundamentals coincide. When $\xi < 1$, the market implied variance of fundamentals is larger than the fundamental one.

We obtain the following comparative statics for yield spreads and expected default rates:

Proposition 2. (*Yield spreads vs. default rates*):

(i) *Yield spreads are always larger than expected default rates: $\Phi(\xi v) > \Phi(v)$, whenever $\xi < 1$.*

(ii) *Yield spreads and default rates are decreasing in credit quality: $\partial\Phi(v)/\partial v > 0$ and $\partial\Phi(\xi v)/\partial v > 0$.*

(iii) *Yield spreads are increasing in information frictions: $\partial\Phi(\xi v)/\partial\xi < 0$.*

This proposition shows that yield spreads exceed expected default rates under noisy information aggregation and limits to arbitrage. Moreover, yield spreads are higher the more severe the information aggregation frictions are. The intuition for these results follows from the insight that information aggregation leads to the market overweighing the tail risks. Since, a priori, default is a tail event under the assumption that $\underline{\theta} < 0$, this immediately implies that the market overweighs default probabilities. This explains why our model delivers yield spreads in excess of default rates, consistent with the credit spread puzzle, and why this effect becomes more pronounced with the severity of information aggregation frictions.

Next, we illustrate that our model is consistent also with the basic comparative statics that were highlighted above: namely, that the puzzle becomes more pronounced for less risky bonds. For this, we consider the comparative statics of the spread ratio $s(\tau, \theta)/d(\tau, \theta)$ with respect to credit quality and information frictions:

Proposition 3. (*Credit spreads ratios*):

(i) **Credit quality:** *The spreads ratio is decreasing in v .*

(ii) **Information frictions:** *The spreads ratio is decreasing in ξ .*

(iii) **Interaction effects:** *The spreads ratio has increasing differences in ξ and v .*

Part (i) shows that the spread ratio is higher for higher quality bonds. Part (ii) shows that the spread ratio is larger when information frictions become more important. Part (iii) establishes an interaction between information frictions and credit quality, by which information frictions have a larger effect on the spread ratio for higher quality bonds.²⁰ The higher the ex ante quality of the bond, the lower is the perceived risk of default. But the less likely a default is, the more the information friction induces the market to overweight the default likelihood. In other words, even if a default may seem very unlikely from an objective point of view, if the market overweighs the tails enough, the market's perception of this risk increases disproportionately relative to the objective likelihood of default.²¹ This proposition thus shows that the comparative statics of our model are consistent with the stylized facts about the credit spread puzzle, which is perhaps the most striking feature of our model.

²⁰In fact, parts (i) and (ii) follow directly from (iii) and the observation that the spread ratio equals 1 for all v , when $\xi = 1$, or for all ξ , when $v = 0$.

²¹While our proof makes use of the tail properties of normal distributions, the result applies whenever the fundamentals are sufficiently thin-tailed.

The next proposition discusses limiting properties of the spread ratio.

Proposition 4. (*Spreads ratio limits*):

(i) *The spreads ratio is unbounded as the default probability approaches 0:*

$$\lim_{v \rightarrow -\infty} \Phi(\xi v) / \Phi(v) = \infty.$$

(ii) *The spreads ratio converges to one, as the default probability approaches 1/2:*

$$\lim_{v \rightarrow 0} \Phi(\xi v) / \Phi(v) = 1.$$

(iii) *The spreads ratio converges to a constant, as information frictions become unboundedly large: $\lim_{\xi \rightarrow 0} \Phi(\xi v) / \Phi(v) = 0.5 / \Phi(v)$.*

Part (i) – (ii) of this proposition show that, for a given level of information frictions, the spreads ratio becomes arbitrarily large, as the bond becomes perfectly safe and approaches unity as ex-ante default risk increases. Thus our model has the potential to account for arbitrarily large spreads ratios, as seen in the data for high quality investment grade bonds. Part (iii) provides a bound on the spread ratios that are feasible, for any given level of credit quality. This bound becomes especially large the safer the bond under consideration, but for a given credit quality it remains finite, and hence provides discipline on the quantitative predictions of our model.

4.2 Persistence and Horizon Effects

We now derive comparative statics with respect to the trading horizon and the persistence of the fundamental process. In particular, we derive closed form solutions for the cases of $\rho = 0$ and $\rho = 1$.

Case with iid Fundamentals. When $\rho = 0$, the market-implied and objective default probabilities are independent of fundamentals. The default events are then iid across periods, and the cumulative default probabilities are given by

$$\Pi(\tau, \theta) = 1 - \left(1 - \Phi\left(\frac{\theta}{\sigma_\theta}\right)\right)^\tau \quad \text{and} \quad \hat{\Pi}(\tau, \theta) = 1 - \left(1 - \Phi\left(\frac{\theta}{\sigma_P}\right)\right)^\tau$$

To a first order, these probabilities are approximated by $\Pi(\tau, \theta) \approx \tau \Phi\left(\frac{\theta}{\sigma_\theta}\right)$ and $\hat{\Pi}(\tau, \theta) \approx \tau \Phi\left(\frac{\theta}{\sigma_P}\right)$ when default risks are small. We then have the following result:

Proposition 5. (*iid Fundamentals*):

When $\rho = 0$, to a first order, the yield spreads, default rates and spreads ratios do not vary with bond maturity and fundamentals.

This proposition shows that effects of bond maturity on yield spreads depend on

persistence in fundamentals.²² Moreover, because fundamentals are fully transitory and the role of current fundamentals for future defaults is nil, spreads are independent of current fundamentals. Thus, if we interpret the fundamental θ as a measure of bond quality, then the impact of bond quality and maturity on yield spreads and expected default rates depends on the persistence of fundamentals.²³

Case with Random Walk Fundamentals. Consider next the case where $\rho = 1$, in which the fundamental follows a random walk. Let $v = (\underline{\theta} - \theta) / \sigma_\theta < 0$ and $\xi = \sigma_\theta / \sigma_P$. As before v measures the financial distress - the closer v is to zero, the higher the default risk - and ξ the severity of information frictions. We reformulate the recursion for Π in terms of u as follows:

$$\Pi(\tau, v) = \Phi(v) + \int_0^\infty \Pi(\tau - 1, v') d\Phi(v' - v)$$

The recursion for $\hat{\Pi}$ in turn can be written as

$$\hat{\Pi}(\tau, v) = \Phi(\xi v) + \int_0^\infty \hat{\Pi}(\tau - 1, v') d\Phi(\xi(v' - v))$$

We have the following lemma:

Lemma 2. *When $\rho = 1$, the market-implied default probability is a rescaling of the objective default probability, $\hat{\Pi}(\tau, v) = \Pi(\tau, \xi v)$, with rescaling factor $\xi = \sigma_\theta / \sigma_P$.*

Thus, in the random walk case, the objective default rates only depend on a normalized distance of fundamentals from the default threshold, and the market-implied default probabilities re-scale this normalized distance by the extent to which the market uses the higher standard deviation of innovations in the fundamental process.

To go further in the characterization of the solution, we proceed with a continuous time approximation of the dynamic trading model. Let Δ denote the length of a time period, let the variance of a one-period innovation to the fundamental be $\sigma_\theta^2 \Delta$, and consider the limit where $\Delta \rightarrow 0$, holding ξ constant. In this limiting case, generations of traders (and their private information) are short-lived, but market noise remains large. Then, the fundamental process $\{\theta_t\}$ converges to a Brownian motion with scale parameter σ_θ^2 and no drift, while the market's standard deviation of fundamentals exceeds the objective standard deviation by a factor of ξ^{-1} , corresponding to a Brownian motion with scale parameter σ_P^2 .

²²While the above result relies on a first-order approximation with small default risk, it holds exactly (i.e. without approximation) in the special case where $c = 0$.

²³Alternatively, we could interpret credit quality as being linked to the threshold $\underline{\theta}$, in which case lower thresholds will translate into a higher spreads ratio in line with the previous results.

In this case, a default occurs the first time that the Brownian motion reaches $\theta_t = \underline{\theta}$, and the default probability is $\Pi(\tau, \theta) = \Pr(\min_{0 \leq t \leq \tau} \theta_t < \underline{\theta} | \theta_0 = \theta)$. At the continuous time limit, $\Pi(\tau, v) = 2\Phi(v/\sqrt{\tau})$ and $\hat{\Pi}(\tau, v) = \Pi(\tau, \xi v) = 2\Phi(\xi v/\sqrt{\tau})$ (Karatzas and Shreve, p. 96, equation (8.4)). Hence, the spread ratio is $\Phi(\xi v/\sqrt{\tau})/\Phi(v/\sqrt{\tau})$. It follows immediately that propositions 2-4 continue to apply identically to the dynamic trading model. In addition, we find the following results:

Proposition 6. (The effect of bond maturity with persistent shocks):

- (i) The spread ratio is decreasing in maturity τ .
- (ii) The spread ratio is unbounded, for a bond with zero maturity in the limit:
 $\lim_{\tau \rightarrow 0} \Phi(\xi v/\sqrt{\tau})/\Phi(v/\sqrt{\tau}) = \infty$.
- (iii) The spread ratio converges to unity, as the horizon grows unboundedly large:
 $\lim_{\tau \rightarrow \infty} \Phi(\xi v/\sqrt{\tau})/\Phi(v/\sqrt{\tau}) = 1$.

In line with our previous results, the reason short-term bonds have a higher spread ratio is that they are safer. Hence, as with credit quality, the above argument about the relative thickness of the tails applies. Conversely, as maturity increases, the ex-ante probability of reaching the default boundary before maturity increases, which reduces the spread ratio.²⁴

5 Quantitative Results

This section describes the calibration of our model and the quantitative results. Section 5.1 discusses the parameter choices for the calibration, section 5.2 presents the quantitative results.

5.1 Calibration

Fundamental process and default thresholds. We interpret one period as a year. The terminal cash value of a zero coupon bond with maturity τ years is a function of the realized state: $p(0, \theta_\tau) = 1$, if $\theta_\tau > \underline{\theta}$, and c otherwise. We then recursively solve for the expected cash value at t periods remaining to maturity, $p(t, \theta_{\tau-t})$, using the recursion

²⁴The continuous time limit serves not just as an approximation, but also as a bound on default probabilities and spread ratios in the discrete time setting. For this observe simply that the discrete time model is based on a default criterion that compares the continuous time fundamental process against the discrete time process only at a discrete, finite set of dates, implying $\Pi(\tau, v) < 2\Phi(v/\sqrt{\tau})$ and $\hat{\Pi}(\tau, v) < 2\Phi(\xi v/\sqrt{\tau})$. At the same time, we can use the probability of default for the final date as a lower bound on the cumulative default probability, so $\Pi(\tau, v) > \Phi(v/\sqrt{\tau})$ and $\hat{\Pi}(\tau, v) > \Phi(\xi v/\sqrt{\tau})$. Combining these inequalities, the actual spreads ratio is bounded by

$$\frac{1}{2} \frac{\Phi(\xi v/\sqrt{\tau})}{\Phi(v/\sqrt{\tau})} < \frac{\hat{\Pi}(\tau, v)}{\Pi(\tau, v)} < 2 \frac{\Phi(\xi v/\sqrt{\tau})}{\Phi(v/\sqrt{\tau})}.$$

in (6). Iterating τ times gives the initial bond price, $p(\tau, \theta)$. Likewise, we compute the cumulative default probabilities as a function of the initial state and the maturity of the bond by applying the recursion in expression (9). As a normalization, we set $\sigma_\theta = 1$ throughout the analysis, and we use the value $\rho = 0.9$ for the persistence of the fundamental process in (1).²⁵ Results and conclusions change very little for alternative values of persistence over a a considerable range (i.e., 0.8 to 0.99, not reported).

To calibrate default thresholds and compute the main statistics for each credit rating category, we proceed as follows. We treat each maturity/credit rating individually. We assume that each bond is issued at an initial state equal to the prior mean: $\theta_0 = 0$. We then calibrate the default threshold $\underline{\theta}(CR, \tau)$ such that the cumulative default probability matches the average default frequencies found in the data. For example, for a Aaa bond with maturity of 10 years, we choose $\underline{\theta}(Aaa, 10)$ such that the cumulative default probability at issuance, conditional on the state $\theta_0 = 0$, equals 0.04%, as reported by Huang and Huang (2012).²⁶ Following the same normalization, we then compute the price of the bond at issuance conditional on the state $\theta_0 = 0$. With this calibration, we need to find 6 default thresholds $\underline{\theta}(CR, \tau)$ for each maturity, where $CR = \{B, Ba, Baa, A, Aa, Aaa\}$. Since our calibration determines the default threshold for each asset class by matching the objective default probabilities, the default thresholds are independent of the degree of informational frictions σ_P .

Informational frictions. While direct measures of informational frictions are not readily available, we calibrate the friction parameter in the model to match the key quantities from the empirical work which studies the impact of analyst forecast dispersion on asset prices. Guntay and Hackbarth (2010) report summary statistics of their data on analyst forecast dispersion about quarterly earnings per share (normalized by the share price), for those firms included in their sample of corporate bonds. They also report similar statistics for the volatility (standard deviation) of earnings.²⁷

We interpret the firm's quarterly earnings per share as a proxy for the fundamental θ . Then D in our model corresponds to the cross-sectional dispersion of fundamentals. It is given by equation (7), and corresponds to the variance of the expectation of θ in

²⁵The normalization of σ_θ is immaterial for the results reported below, as the only statistic that matters for spreads (for a given maturity and credit rating, given that we match default probabilities) is the ratio σ_P/σ_θ .

²⁶To allow a more direct comparison between the explanatory power of our model and the battery of models included in Huang and Huang (2012), we calibrate CDP's of 4 and 10 year bonds to fit the empirical averages reported by Moody's for the 1970-1998 period. The magnitude of the quantitative results is practically unchanged if we use the longer sample ending in 2007.

²⁷In Guntay and Hackbarth (2010), volatility of earnings is the standard deviation of the last 8 quarterly earnings reported by each firm.

the cross-section of informed traders:

$$D = \int \{E[\theta_{+1}|x, z] - \int E[\theta_{+1}|x, z]dx\}^2 dx.$$

Taking account of the normalization of fundamental volatility σ_θ to 1, we thus calibrate the square root of belief dispersion \sqrt{D} , in our model to be equal to the ratio between the square root of the cross sectional dispersion of analyst's forecasts, and the measured volatility of quarterly earnings.

Table 4: Range of σ_P values

σ_u	1.0	1.5	2.0	2.5	3.0	3.5	4.0
σ_P							
Aaa							
d/ σ_θ = mean	1.03	1.05	1.08	1.12	1.16	1.21	1.26
d/ σ_θ = + 1/2 st. dev.	1.06	1.10	1.15	1.21	1.28	1.36	1.44
Aa							
d/ σ_θ = mean	1.08	1.12	1.18	1.26	1.34	1.44	1.54
d/ σ_θ = + 1/2 st. dev.	n.a.	n.a.	1.37	1.50	1.65	1.82	1.99
A							
d/ σ_θ = mean	1.07	1.11	1.17	1.24	1.32	1.41	1.51
d/ σ_θ = + 1/2 st. dev.	n.a. ¹	1.23	1.34	1.47	1.61	1.76	1.92
Baa							
d/ σ_θ = mean	1.08	1.12	1.19	1.26	1.35	1.44	1.55
d/ σ_θ = + 1/2 st. dev.	n.a.	1.25	1.36	1.49	1.64	1.80	1.97
Ba							
d/ σ_θ = mean	1.08	1.12	1.18	1.26	1.34	1.44	1.54
d/ σ_θ = + 1/2 st. dev.	n.a.	n.a.	1.38	1.51	1.67	1.83	2.01
B							
d/ σ_θ = mean	1.07	1.11	1.16	1.23	1.31	1.39	1.49
d/ σ_θ = + 1/2 st. dev.	n.a.	1.21	1.31	1.43	1.56	1.70	1.85

¹ n.a.: a real solution for β does not exist. ² d: square root of dispersion, D .

For each measure of dispersion D , we present several values of noise trading shock volatility, σ_u , and for each we compute the values of β that satisfy that value of D .²⁸ The resulting levels of σ_P for each value of σ_u are reported in Table 4, which displays the values of the informational friction parameter that obtains using the mean of dispersion, as well as the mean plus 1/2 its standard deviation. All the statistics for the cross-sectional analyst forecast dispersion, $d = \sqrt{D}$, and for the volatility of earnings, σ_θ , were

²⁸Since equation (7) represents a quadratic equation on β , whenever it has a real solution, there are two values of β that satisfy the relation. These are not reported but available upon request.

provided by Guntay and Hackbarth (2010) at the credit-rating level.²⁹ For example, if the volatility of noise trading shocks is 1 (first column), then the normalized analyst forecast dispersion d/σ_θ evaluated at the sample mean for Aaa firms (first row) implies a value of $\sigma_P = 1.03$. If we keep the same value for σ_u , but use for d/σ_θ the mean plus 0.5 times the standard deviation for Aaa-rated firms (second row), we obtain $\sigma_P = 1.06$. Notice however that these magnitudes increase significantly as we move to higher levels of noise trading volatility. For a value of $\sigma_u = 2$, σ_P reaches 1.15 for Aaa bonds, when evaluated at d/σ_θ equal the sample mean plus half a st. dev.

In lower rating categories, the implied values for σ_P for a given level of noise trading volatility also increase, as forecast dispersion is generally lower in the data for Aaa firms. For Aa bonds in firms with d/σ_θ equal to the sample mean, σ_P equals 1.08 if noise trading volatility is $\sigma_u = 1$, and increases up to $\sigma_P = 1.34$ for $\sigma_u = 3$. For some combination of dispersion and noise trading volatility however, there is no value for the private signal precision (β) that satisfies the quadratic equation for dispersion in (7). Simply put, for a given level of earnings volatility σ_θ and noise trading σ_u , the magnitude of forecast dispersion consistent with the model is bounded. Therefore, too large a value for the observed level of dispersion D may not be consistent with certain values of σ_u . For instance, all rating categories except Aaa have levels of dispersion that are too high to be generated by a noise trading volatility as low as $\sigma_u = 1$, when dispersion is evaluated at the sample mean plus half a standard deviation. To be consistent with the model, the level of noise trading volatility must be larger than 1, and larger even than 1.5 in some cases. Intuitively, for a given level of fundamental volatility, the market signal z must be noisy enough to keep informed trader's forecast sufficiently dispersed. For this to occur, noise trading volatility must be large enough.

We interpret this as evidence that the relevant range of noise trading volatility in the data, although not directly observed, must be large enough to be consistent with the observed levels of forecast dispersion, which in turn is consistent with relatively large values of σ_P in Table 1, in the range of at least 1.3, for all bond categories except Aaa.

5.2 Calculated Credit Spreads

5.2.1 Spread ratios

We now discuss our results for the spreads ratio of market yield spreads to expected loss rates in our model. The top, middle, and bottom panels of Table 5 show the simulated spread ratios that obtain in our simulated model for maturities of 4, 10 and 20 years,

²⁹We thank the authors for generously sharing the statistics of this ratio, which is not reported in their paper. The figures used here correspond to the sub-sample 1987-1998.

for different values of σ_P and credit ratings.

In line with statement i) in Proposition 3, the spread ratios (for any value of $\sigma_P > 1$) decrease monotonically in credit quality. Indeed, highly rated bonds have an ex-ante very small probability of default. For these bonds, the overweighting of the tails of the fundamental distribution generated by the informational frictions has a much stronger impact on spreads, relative to actual default probabilities, than in low quality bonds. Intuitively, if a bond is very risky ex-ante (i.e., with an ex-ante default probability close to 1/2), then overweighting the tails of the fundamental distribution has virtually no impact on spreads. In contrast, the price of a bond that defaults whenever the fundamentals falls, say, 3 standard deviations below the mean, is extremely sensitive to how the tails of the distribution are weighted in the market. For example, the spread ratio of a 4-year Aaa bond when $\sigma_P = 1.3$ equals 14.1, that is, the market-implied spread is about 14 times as large as the spread which would obtain in the absence of information frictions. For the same degree of information frictions, this ratio is only 4, 2.1, and 1.5 for Baa, Ba and B bonds, respectively.

Consistent with part ii) of Proposition 3, Table 5 shows how higher information frictions increase spread ratios monotonically, for any credit rating. This is to be expected from our previous discussion, as higher σ_P implies the pricing of a security in the market assigns more weight to tail events. This higher weighting increases the spreads for all bonds with ex-ante probability of default less than 50% per year (a condition satisfied by all bonds under study). For example, the spread ratio for 4-year Aaa bonds is 4.8 when $\sigma_P = 1.15$, but increase up to 30.8 for values of σ_P as large as 1.45.

The interaction effect of credit quality and informational frictions predicted by part iii) of Proposition 3 is quantitatively significant. The increase in spread ratios as we move to higher degrees of informational frictions is particularly large for high quality bonds. For instance, at 4 year maturities, an increase in σ_P from 1.15 to 1.45 raises the spread ratio in Aaa bonds by a factor of 6.5 times. The same increase in frictions raises spread ratios by only 4.2 and 3.7 times for Aa and A bonds, respectively. Towards the lower end of credit quality, the effect is even less pronounced. The same increase in frictions leads to spread ratios which are roughly 2.7, 1.7, and 1.3 times as large for Baa, Ba, and B bonds.

Next, we discuss how bond maturity affects spread ratios. Proposition 5 states that spread ratios should be independent of maturity when fundamentals (and hence default probabilities) are iid across time. On the other hand, Proposition 6 states that maturity should decrease spread ratios monotonically when fundamentals follow a random walk. This follows from the fact that in such case, an increase in maturity increases CDP's more than proportionally - an increase in maturity is in fact isomorphic to a decline in

credit quality for a given maturity. Hence a longer term bond is perceived as less safe, implying that the spreads ratio declines with maturity.

Table 5: Spread ratios

Data ¹		σ_P						
		1.00	1.15	1.30	1.45	1.60	1.75	1.90
4 yr								
Aaa	93.9	1.0	4.8	14.1	30.8	55.7	88.1	126.9
Aa	19.9	1.0	3.3	7.4	13.6	21.4	30.6	40.6
A	20.3	1.0	3.0	6.4	11.1	16.9	23.5	30.6
Baa	9.8	1.0	2.3	4.0	6.2	8.5	11.0	13.5
Ba	2.9	1.0	1.5	2.1	2.6	3.1	3.5	4.0
B	1.6	1.0	1.3	1.5	1.7	1.8	2.0	2.1
10 yr								
Aaa	12.4	1.0	2.8	5.7	9.6	14.2	19.1	24.3
Aa	14.2	1.0	2.6	5.2	8.4	12.2	16.2	20.3
A	12.6	1.0	2.4	4.4	6.8	9.5	12.3	15.1
Baa	6.9	1.0	1.9	3.0	4.1	5.3	6.4	7.5
Ba	2.9	1.0	1.4	1.7	2.0	2.3	2.5	2.7
B	1.9	1.0	1.2	1.3	1.4	1.5	1.6	1.6
20 yr								
Aaa	11.7	1.0	2.5	4.7	7.4	10.4	13.4	16.3
Aa	15.2	1.0	2.5	4.5	7.0	9.7	12.4	15.1
A	10.6	1.0	2.1	3.5	5.0	6.6	8.1	9.5
Baa	7.2	1.0	1.7	2.4	3.2	3.9	4.5	5.1
Ba		1.0	1.3	1.5	1.7	1.8	2.0	2.1
B		1.0	1.2	1.3	1.4	1.5	1.5	1.6

¹ The empirical ratios are from table 1.

Comparing the different panels of Table 5 reveals that the prediction from the fully persistent case holds in all cases under consideration. For example, 4-year Aaa bonds exhibit larger spread ratios than 10-year Aaa bonds –14.1 vs. 5.7, for an informational friction parameter $\sigma_P = 1.3$. For the same level of informational frictions, increasing maturity to 20 years further decreases the spread ratio –from 5.7 to 4.7. This is to be expected, as an increase in maturity of 2.5 times (i.e., from 4 to 10 years) increases cumulative default probabilities by a factor of almost 18 times. Because the 10 year Aaa bond is then riskier (in terms of per year default probability) than the 4 year bond, its pricing is less sensitive to overweighting of the distribution tails, and hence have lower spread ratios. Likewise, the 20 year Aaa has a cumulative default probability of nearly 2.7 times as high as the 10 year bond, so it is riskier as well in per-annum terms, and

hence shows lower spread ratios. This conclusion holds true for all bond categories and all maturities, given the cumulative default probabilities and yield spreads observed in the data.

5.2.2 Absolute spreads

Table 6 presents the absolute credit spreads for 4, 10, and 20 year bonds that we obtain for different values of the informational frictions parameter, σ_P . The columns on the left, show the absolute spreads generated by the model for each value of the informational friction parameter, σ_P ; the columns on the right present the calculated spreads as a fraction (in %) of the average spreads observed in the data.

The upper part of the table presents the results for 4-year bonds. The first column in the top-left table (calculated spreads), shows the spreads generated by the model when there are no informational frictions, $\sigma_P = 1$. For Aaa securities, the calculated spread is 0.51 bp, which accounts for only 1.1% of the empirical value reported in Table 1. As we increase information frictions, the calculated spread increases substantially. For $\sigma_P = 1.3$, for instance, it generates 7.2 bp, or about 15.8% of the empirical counterpart, and almost 35% of the observed spread for $\sigma_P = 1.45$. As a reference point, to account for the whole historical spread of the Aaa, 4 year bond, the model would need an informational frictions value $\sigma_P = 1.75$. As we lower credit quality, the model gets closer to the empirical spreads for smaller values of σ_P . For Aa bonds, a value of $\sigma_P = 1.45$ explains almost 69% of the data, needing less than $\sigma_P = 1.6$ to account for the whole spread.

To put these results in perspective, note that a battery of structural models developed in Huang and Huang (2012) explain a very small portion of the spreads of high investment grade bonds—around 2, 9, and 10% for 4-year Aaa, Aa, and A categories, respectively. The value of the spreads generated by our model assuming $\sigma_P = 1.3$ deliver significantly higher numbers, especially for the the bonds of Aaa and A categories: at ratios near 16%, 38%, and 32% of the empirical values. Our discussion of the calibration of the informational frictions parameter in Table 4 suggests that $\sigma_P = 1.3$ is consistent with the the joint observations of forecast dispersion and quarterly earnings volatility reported by Guntay and Hackbarth (2010), in particular for the Aa and A categories.

Table 6: Calculated vs. observed yield spreads

Data ¹		Calculated spreads: $\sigma_P =$							Fraction explained* (%): $\sigma_P =$						
		1.00	1.15	1.30	1.45	1.60	1.75	1.90	1.00	1.15	1.30	1.45	1.60	1.75	1.90
4 yr															
Aaa	46	0.5	2.5	7.2	15.9	28.7	45.4	65.4	1.1	5.3	15.8	34.6	62.4	98.8	142.3
Aa	56	2.8	9.3	21.2	38.6	60.9	86.9	115.5	5.1	16.6	37.8	69.0	108.8	155.2	206.3
A	87	4.4	13.0	27.8	48.4	73.8	102.7	133.8	5.0	14.9	31.9	55.7	84.9	118.0	153.7
Baa	149	15.4	35.0	62.1	94.9	131.2	169.2	207.8	10.4	23.5	41.7	63.7	88.0	113.6	139.5
Ba	310	107.5	163.9	221.8	278.1	331.4	380.9	426.4	34.7	52.9	71.5	89.7	106.9	122.9	137.6
B	470	306.1	383.1	450.8	509.7	560.9	605.4	644.2	65.1	81.5	95.9	108.4	119.3	128.8	137.1
10 yr															
Aaa	47	3.8	10.7	21.8	36.5	53.9	72.7	92.4	8.1	22.7	46.4	77.7	114.6	154.8	196.5
Aa	69	5.0	13.1	25.7	41.9	60.4	80.3	100.7	7.2	19.0	37.3	60.7	87.6	116.4	145.9
A	96	7.7	18.4	33.8	52.4	73.1	94.6	116.2	8.0	19.2	35.2	54.6	76.1	98.5	121.0
Baa	150	22.2	42.2	65.9	91.3	117.1	142.1	166.0	14.8	28.1	43.9	60.9	78.0	94.7	110.6
Ba	310	107.8	146.7	182.5	214.8	243.4	268.7	290.9	34.8	47.3	58.9	69.3	78.5	86.7	93.9
B	470	242.1	281.1	313.1	339.3	361.2	379.4	394.9	51.5	59.8	66.6	72.2	76.8	80.7	84.0
20 yr															
Aaa	59	5.1	12.8	24.0	37.6	52.5	67.8	82.9	8.6	21.7	40.7	63.7	89.0	115.0	140.5
Aa	87	5.7	13.9	25.7	39.8	55.0	70.6	85.8	6.5	16.0	29.5	45.7	63.3	81.1	98.6
A	117	11.0	23.1	38.4	55.2	72.3	89.0	104.8	9.4	19.8	32.8	47.2	61.8	76.1	89.6
Baa	198	27.3	46.4	66.9	86.9	105.8	123.1	138.7	13.8	23.4	33.8	43.9	53.4	62.2	70.1
Ba		100.1	127.0	149.9	169.1	185.1	198.5	209.9	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
B		150.1	174.4	193.6	208.9	221.3	231.4	239.8	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

¹ Yield Spreads for Aaa-Baa are from Duffee (1998); Ba and B from Caouette, Altman, and Narayanan (1998).

We now turn to the results for 10-year bonds, which are displayed in the middle of table 6. If there are no informational frictions ($\sigma_P = 1$), then as before the spreads delivered by the model are low relative to the data. For Aaa instruments, the risk-neutral model generates a spread of 3.8 bp, which amounts to slightly over 8% of the empirical spreads. Similarly, Aa and A bonds have calculated spreads which explain only 7.2 and 8% of the data, respectively. For the lower rated bonds, the model with no frictions explains about 15, 35, and 52% of the observed spreads for the Baa, Ba, and B credit ratings, respectively. As we increase σ_P , the model can account for a substantially higher fraction of spreads, across all categories. For $\sigma_P = 1.3$, for instance, the model explains over 46, 37, and 35% of the data for high investment grade (Aaa, Aa, and A, respectively), and 44, 59 and 67% for the riskier securities (Baa, Ba, and B, respectively). Notice once again how increases in informational frictions have a particularly large effects in the explanatory power of the model for the high investment grade securities, an issue we will discuss in more detail momentarily. Compared to the structural models in Huang and Huang (2012), these figures are significantly higher than the 16, 16 and 19% of spreads for Aaa, Aa, and A, respectively.

The lower part of the table presents the results for 20 year bonds. A similar pattern of credit spreads emerges: as we increase informational frictions, the model delivers spreads which account for an increasing fraction of the empirical spread observations. For $\sigma_P = 1.3$, it explains about 41, 30 and 33% for Aaa, Aa and A categories.³⁰

The impact of maturity on spreads is apparent from the comparison of credit spreads for a given credit rating across panels. The effect of horizons on absolute spreads depends on the bond rating. Consider the case of Aaa bonds. At 4 years maturity, Aaa bonds have a CDP of 0.04%. Increasing maturity 2.5 times (to 10 years) increases the CDP to 0.77% (more than 19 times). Not surprisingly then, the yearly spread in the data is higher for the 10 year, Aaa bond, for all values of the informational frictions parameter considered in the Table. If we now compare 10 to 20 years, we see that the CDP for Aaa bonds increases to 2.05%, which is more than twice 0.77%. Hence, fundamental risk (per year) is slightly higher for 20 year bonds. When the informational frictions parameter is at the lower end – say, σ_P less than 1.45 –, the dominant force is fundamental risk, and consequently the doubling of maturity increases the yearly spread. However, if information frictions are sufficiently large, the fact that safer securities have higher *spread ratios* (i.e., the increasing differences, or interaction effect between credit quality and informational frictions highlighted in Proposition 3,

³⁰If one is willing to accept larger values of σ_P , in the range of 1.6 to 1.9, the fraction of spreads our model is able to account for is quickly increasing, to the point where we are able to account for them entirely. This interpretation of the model, however, would seem inconsistent with the reduced-form evidence in Guntay and Hackbarth (2010) and Buraschi, Trojani, and Vedolin (2008) attributing only a fraction of the spreads to disagreement among forecasters.

part iii)) becomes the dominant force, implying *lower* absolute spreads for the 20 year instruments.

For some bonds in the sample – notably the Ba and B categories –, increasing maturities from 10 to 20 years actually decreases the yearly default risk (for instance, the CDP of 20 year B bonds is 52.9%, less than twice the CDP of 10 year bonds, at 43.9%). For these securities, the fundamental risk effect dominates, and our model generates higher absolute spreads for the 10 year than for 20 year securities, for all values of the informational friction parameter.³¹

Finally, our model can also account for a sizeable fraction of the Baa-Aaa spread. This spread is one of the main puzzles of bond pricing data, as it is unlikely to be determined by tax asymmetries, callability features, or liquidity premia, which should be roughly similar across investment grade securities (see Chen et al., 2009). In particular, of the 103 bp spread in the data³² for 4 year bonds, our model explains 55 bp, or 53%. For 10 year bonds, the spread is also 103 bp, of which our model explains 44 bp, or 43%. For 20 year bonds, the model accounts for about 43 bp, roughly 31% of the 139 bp empirical spread.

To summarize, this section has documented that heterogenous information can explain a sizable fraction of the credit spread puzzle, in particular for high investment grade, short-term securities. Under a reasonable parameterization for the degree of informational frictions (i.e., leading to a value of $\sigma_P = 1.3$) our model can explain close to 16% of observed spreads for Aaa, 4 year bonds, which present the most challenging evidence of credit spreads through the lens of no-arbitrage pricing. For 10 and 20 year Aaa, as well as for most of the rest of investment grade bonds of 4, 10, and 20 year maturities, our model explains over 30% of observed spreads. These figures are large in comparison to what most structural models in the current literature deliver.

6 Concluding Remarks

In this paper we developed a theory of corporate bond pricing based on heterogeneous information and limits to arbitrage. Our model offers pricing implications of noisy information aggregation for a class of securities whose payoffs are non-linear functions underlying fundamentals, while retaining a tractable and parsimonious structure.

³¹This result may also be a consequence of survivor bias in the data, or mean reversion in the model. Survivor bias (in the data) arises if for a given ratings category, the worse companies are more likely to default, and thus over time the pool shifts towards higher quality survivors. From an ex ante perspective this makes the 20-year bonds seem safer than 10-year bonds of equal intrinsic quality, and may account for the lower long-term spreads. Mean reversion in fundamentals implies that forecasts of default sufficiently far into the future should be virtually independent of initial conditions. But this brings us closer to the characterization offered by the iid case, in which maturity length has no effect on spreads.

³²Spreads corresponds to those reported by Duffee (1998).

Quantitatively, the model explains a sizable portion of the spreads, and compares favorably with respect to other asset pricing models that place significantly more structure on the environment. In particular our theory generates the largest relative departures between market spreads and underlying default risk for high investment grade, short maturity instruments. This is in line with the empirical evidence about the credit spread puzzle, and represents the major challenge in fitting the data for most existing structural models in the literature.

References

- [1] Albagli, Elias, Christian Hellwig, and Aleh Tsyvinski (2012). “A Theory of Asset Prices based on Heterogeneous Information.” NBER Working paper No. 17548.
- [2] Bacchetta, Philippe, and Eric van Wincoop (2006). “Can Information Heterogeneity Explain the Exchange Rate Determination Puzzle?” *American Economic Review*, 96 (3), 552-576.
- [3] Banerjee, Snehal (2011). “Learning from Prices and the Dispersion in Beliefs.” *Review of Financial Studies*, 24 (9), 3025-68.
- [4] Barlevy, Gadi, and Pietro Veronesi (2000). “Information Acquisition in Financial Markets.” *The Review of Economic Studies*, 67 (1), 79-90.
- [5] _____ (2003). “Rational Panics and Stock Market crashes.” *Journal of Economic Theory*, 110 (2), 234-263.
- [6] Bhamra, Harjoat S., Lars-Alexander Kuehn, and Ilya A. Strebulaev (2010). “The Aggregate Dynamics of Capital Structure and Macroeconomic Risk.” *Review of Financial Studies*, 23 (12), 4187-4241.
- [7] Biais, Bruno, Peter Bossaerts, and Chester Spatt (2010). “Equilibrium Asset Pricing and Portfolio Choice under Asymmetric Information.” *Review of Financial Studies*, 23 (4), 1503-1543.
- [8] Breon-Drish, Bradyn (2012). “On Existence and Uniqueness of Equilibrium in Noisy Rational Expectations Economies.” Working paper, Stanford University.
- [9] Buraschi, Andrea, Fabio Trojani, and Andrea Vedolin (2008). “Economic Uncertainty, Disagreement, and Credit Markets.” Working paper.
- [10] Caouette, John B., Edward I. Altman, and Paul Narayanann (1998). *Managing credit risk: the next great financial challenge*. Vol. 2. John Wiley & Sons.
- [11] Chen, Long, Pierre Collin-Dufresne, and Robert S. Goldstein (2009). “On the Relation Between the Credit Spread Puzzle and the Equity Premium Puzzle.” *Review of Financial Studies*, 22 (9), 3367-3409.
- [12] Collin-Dufresne, Pierre, Robert S. Goldstein, and J. Spencer Martin (2001). “The Determinants of Credit Spread Changes.” *The Journal of Finance*, 56 (6), 2177-2207.

- [13] Diamond, Douglas W., and Robert E. Verrecchia (1981). “Information Aggregation in a Noisy Rational Expectations Economy.” *Journal of Financial Economics*, 9 (3), 221-235.
- [14] Diether, Karl B., Christopher J. Malloy, and Anna Scherbina (2002). “Differences of Opinion and the Cross Section of Stock Returns.” *The Journal of Finance*, 57 (5), 2113-2141.
- [15] Duffee, Gregory (1998). “Treasury Yields and Corporate Bond Yield Spreads: An Empirical Analysis.” *Journal of Finance*, 53 (6), 2225-2242.
- [16] Duffie, Darrell, and David Lando (2001). “Term Structures of Credit Spreads with Incomplete Accounting Information.” *Econometrica*, 69 (3), 633-664.
- [17] Elton, Edwin, Martin J. Gruber, Deepak Agrawal, and Christopher Mann (2001). “Explaining the Rate Spread on Corporate Bonds.” *The Journal of Finance*, 56 (1), 247-277.
- [18] Gabaix, Xavier (2012). “Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance.” *Quarterly Journal of Economics*, 127 (2), 645-700.
- [19] Giesecke, Kay, Francis A. Longstaff, Stephen Schaefer, and Ilya Strebulaev (2011). “Corporate Bond Default Risk: A 150-year Perspective.” *Journal of Financial Economics*, 102 (2), 233-250.
- [20] Gomes, João, and Lukas Schmid (2013). “Equilibrium Credit Spreads and the Macroeconomy.” Working paper.
- [21] Gourio, François (2013). “Credit Risk and Disaster Risk.” *AEJ: Macroeconomics*, 5 (3), 1-34.
- [22] Gromb, Denis, and Dimitri Vayanos (2010). “Limits of Arbitrage.” *Annual Review of Financial Economics*, 2, 251-275.
- [23] Grossman, Sanford J. and Joseph E. Stiglitz (1980). “On the Impossibility of Informationally Efficient Markets.” *American Economic Review*, 70 (3), 393-408.
- [24] Güntay, Levent, and Dirk Hackbarth (2010). “Corporate Bond Credit Spreads and Forecast Dispersion.” *Journal of Banking & Finance*, 34 (10), 2328-2345.
- [25] Harrison, Michael and David Kreps (1978). “Speculative Investors in a Stock Market with Heterogeneous Expectations.” *Quarterly Journal of Economics*, 92 (2), 323-336.

- [26] He, Hua and Jiang Wang (1995). "Differential Information and Dynamic Behavior of Stock Trading Volume." *Review of Financial Studies*, 8 (4), 919-972.
- [27] Hellwig, Martin (1980). "On the Aggregation of Information in Competitive Markets," *Journal of Economic Theory*, 22, 477-498.
- [28] Hong, Harrison, and David Sraer (2013). "Quiet Bubbles," *Journal of Financial Economics*, forthcoming.
- [29] Huang, Jing-zhi, and Ming Huang (2012). "How much of the Corporate-Treasury Yield Spread is due to Credit Risk?." *Review of Asset Pricing Studies*, 2 (2), 153-202.
- [30] Johnson, Timothy C. (2004). "Forecast Dispersion and the Cross Section of Expected Returns." *The Journal of Finance*, 59 (5), 1957-1978.
- [31] Keenan, Sean, Igor Shtogrin, and Jorge Sobehart (1998). "Historical Default Rates of Corporate Bond Issuers, 1920 - 1998". Moody's special comment report.
- [32] Karatzas, Ioannis Autor. *Brownian motion and stochastic calculus*. Ed. Steven Eugene Shreve. Vol. 113. Springer, 1991.
- [33] McQuade, Timothy (2013). "Stochastic Volatility and Asset Pricing Puzzles". Working paper.
- [34] Miller, Edward M. (1977). "Risk, Uncertainty, and Divergence of Opinion." *The Journal of Finance*, 32 (4), 1151-1168.
- [35] Scheinkman, Jose and Wei Xiong (2003). "Overconfidence and Speculative Bubbles." *Journal of Political Economy*, 111 (6), 1183-1220.
- [36] Townsend, Robert (1983). "Forecasting the Forecasts of Others." *Journal of Political Economy*, 91 (4), 546-588.
- [37] Vanden, Joel M. (2008). "Information Quality and Options." *Review of Financial Studies*, 21(6), 2635-2676.
- [38] Veldkamp, Laura (2011). *Information Choice in Macroeconomics and Finance*. Princeton University Press.
- [39] Vives, Xavier (2008). *Information and Learning in Markets*. Princeton University Press.
- [40] Wang, Jiang (1993). "A Model of Intertemporal Asset Prices Under Asymmetric Information." *Review of Economic Studies*, 60 (2), 249-282.

- [41] Wang, Jiang (1994). “A Model of Competitive Stock Trading Volume.” The Journal of Political Economy, 102 (1), 127-168.
- [42] Watanabe, Masahiro (2008). “Price Volatility and Investor Behavior in an Overlapping Generations Model with Information Asymmetry.” Journal of Finance, 63 (1), 229-272.
- [43] Yu, Fan (2005). “Accounting Transparency and the Term Structure of Credit Spreads.” Journal of Financial Economics, 75 (1), 53-84.

7 Appendix

Proof of Proposition 1:

Proof. The above argument is complete once we have shown that $p(\tau, \theta)$ is monotonic in θ , for all τ . To show this, we proceed by induction. When $\tau = 0$, $p(0, \theta)$ is non-decreasing in θ , with a strict upwards jump at $\underline{\theta}$, and moreover $p(0, \theta) \geq 0$ for all θ . Suppose therefore that $p(\tau, \theta_{T-\tau})$ is non-decreasing in $\theta_{T-\tau}$ and $p(\tau, \theta_{T-\tau}) \geq c$ for all $\theta_{T-\tau} \geq \underline{\theta}$. By virtue of the conditional normal distributions, $\theta_{+1}|x, \theta, z$ is first-order stochastically increasing in all its arguments, and $z|\theta$ and is first-order stochastically increasing in θ , so $\mathbb{E}(p(\tau - 1, \theta_{+1})|x = z, \theta, z)$ is monotonically increasing in θ , and $\mathbb{E}(p(\tau - 1, \theta_{+1})|x = z, \theta, z) \geq c$, by virtue of the induction hypothesis. Therefore, $p(\tau, \theta)$ is monotonically increasing in θ , with an upwards discontinuity at $\underline{\theta}$, which completes the result. \square

Proof of Lemma 1:

Proof. The prior of θ_{t+1} is given by $\mathcal{N}(\rho\theta_t, \sigma_\theta^2)$, and the conditional distribution of x_t , and z_t on θ_{t+1} are given by $\mathcal{N}(\theta_{t+1}, \beta^{-1})$, and $\mathcal{N}(\theta_{t+1}, \frac{\sigma_u^2}{\beta})$ respectively. Applying the Bayesian updating formula on Gaussian random variables, $\theta_{t+1}|\theta_t, x_t = z_t, z_t$ has distribution:

$$\mathcal{N}\left(\frac{\rho\theta_t/\sigma_\theta^2 + (\beta + \beta/\sigma_u^2)z_t}{1/\sigma_\theta^2 + \beta + \beta/\sigma_u^2}, \frac{1}{1/\sigma_\theta^2 + \beta + \beta/\sigma_u^2}\right)$$

Since $z_t|\theta_t$ has distribution $\mathcal{N}(\rho\theta_t, \sigma_\theta^2 + \sigma_u^2\beta^{-1})$, $\theta_{t+1}|\theta_t$ is distributed normally with mean $\rho\theta_t$ and variance:

$$\frac{1}{1/\sigma_\theta^2 + \beta + \beta/\sigma_u^2} + \left(\frac{\beta + \beta/\sigma_u^2}{1/\sigma_\theta^2 + \beta + \beta/\sigma_u^2}\right)^2 \left(\sigma_\theta^2 + \frac{\sigma_u^2}{\beta}\right)$$

With some algebra, this reduces to $Var(\theta_{t+1}|\theta_t) = \sigma_\theta^2 + (1 + \sigma_u^2)D$, where

$$D \equiv \frac{\beta}{(1/\sigma_\theta^2 + \beta + \beta\sigma_u^2)^2}.$$

□

Proof of Proposition 2:

Proof. Since $\xi < 1$ and $v < 0$, then $\xi v > v$, which implies $\Phi(\xi v) > \Phi(v)$. Note also that both are increasing in v , and since $v < 0$, ξv is decreasing in ξ , so the market-implied spread yield, $\Phi(\xi v)$, is also decreasing with ξ . □

Proof of Proposition 3:

Proof. The spread ratio is defined as $\Phi(\xi v)/\Phi(v)$, where $v < 0$ and $\xi \in (0, 1]$. Part (ii) is then immediate for any $v < 0$.

For parts (i) and (iii), take the first order derivative with respect to v , the ratio is decreasing if and only if:

$$\frac{\xi\phi(\xi v)}{\Phi(\xi v)} \leq \frac{\phi(v)}{\Phi(v)}.$$

Now, notice that this inequality is automatically satisfied as an equality when $\xi = 1$. Therefore, parts (i) and (iii) follow immediately, if we can show that $\frac{\xi\phi(\xi v)}{\Phi(\xi v)}$ is increasing in ξ . To this end notice that

$$\begin{aligned} \frac{d}{d\xi} \left[\frac{\xi\phi(\xi v)}{\Phi(\xi v)} \right] &= \frac{[\phi(\xi v) - (\xi v)^2\phi(\xi v)]\Phi(\xi v) - \xi v\phi(\xi v)^2}{\Phi(\xi v)^2} \\ &= \frac{\phi(\xi v)}{\Phi(\xi v)} \left\{ 1 - (\xi v) \left((\xi v) + \frac{\phi(\xi v)}{\Phi(\xi v)} \right) \right\} \end{aligned}$$

Since $-(\xi v) > 0$ and

$$\xi v + \frac{\phi(\xi v)}{\Phi(\xi v)} = \frac{\int_{-\infty}^{\xi v} (\xi v - h)\phi(h)dh}{\Phi(\xi v)} \geq 0,$$

it follows that $\frac{d}{d\xi} \left[\frac{\xi\phi(\xi v)}{\Phi(\xi v)} \right] > 0$. □

Proof of Proposition 4:

Proof. For (i), we have from L'Hopital Rule:

$$\lim_{v \rightarrow -\infty} \frac{\Phi(\xi v)}{\Phi(v)} = \lim_{v \rightarrow -\infty} \frac{\xi\phi(\xi v)}{\phi(v)} = \lim_{v \rightarrow -\infty} \xi \exp\left\{\frac{v^2}{2}(1 - \xi^2)\right\} = \infty$$

The other limits are obvious. \square

Proof of Proposition 5:

Proof. Using the Taylor's expansion on the expression for $\Pi(\tau, \theta)$ and $\hat{\Pi}(\tau, \theta)$ and ignoring higher order terms, we get $\Pi(\tau, \theta) = \tau\Phi(\frac{\theta}{\sigma_\theta})$ and $\hat{\Pi}(\tau, \theta) = \tau\Phi(\frac{\theta}{\sigma_P})$, which implies $\Pi(\tau, \theta)/\hat{\Pi}(\tau, \theta) = \Phi(\frac{\theta}{\sigma_\theta})/\Phi(\frac{\theta}{\sigma_P})$. So the spread ratio is independent of τ and θ . \square

Proof of Lemma 2:

Proof. We prove by induction. When $\tau = 1$, $\Pi(1, v) = \Phi(v)$ and $\hat{\Pi}(1, v) = \Phi(\xi v)$, which verifies the statement in the Lemma. Moreover, suppose that $\hat{\Pi}(\tau, v) = \Pi(\tau, \xi v)$, then:

$$\begin{aligned} \hat{\Pi}(\tau + 1, v) - \Pi(\tau + 1, \xi v) &= \int_0^\infty \hat{\Pi}(\tau, v') d\Phi(\xi(v' - v)) - \int_0^\infty \Pi(\tau, v') d\Phi(v' - \xi v) \\ &= \int_{-\xi v}^\infty \hat{\Pi}(\tau, v + \frac{1}{\xi}h) d\Phi(h) - \int_{-\xi v}^\infty \Pi(\tau, \xi v + h) d\Phi(h) \\ &= 0, \end{aligned}$$

where the last step makes use of the induction hypothesis. \square

Proof of Proposition 6:

Proof. Follows directly from the comparative statics w.r.t. v in propositions 2-4. \square